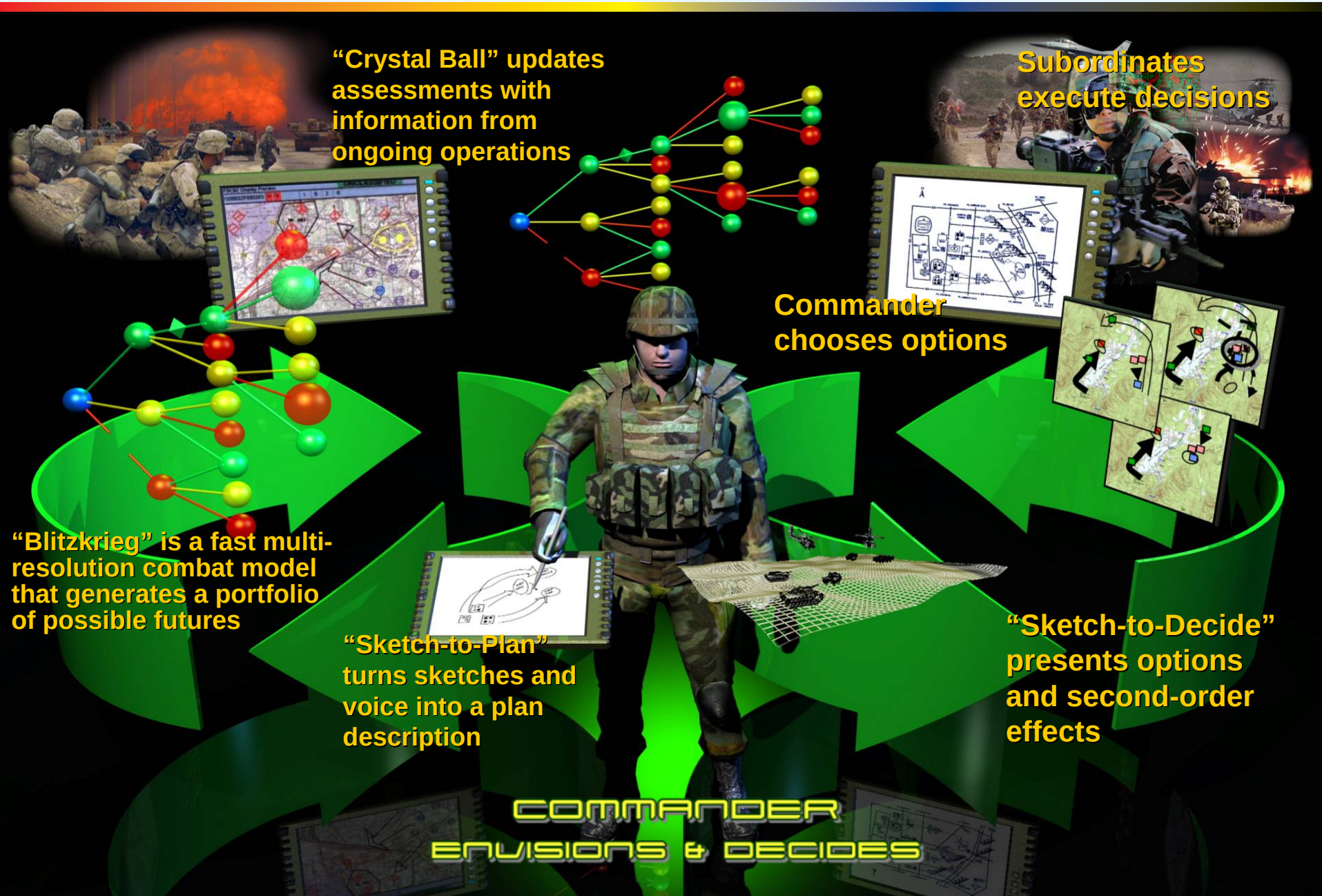
DEEP GREEN

Industry Day

COL John "Buck" Surdu, PM
Information Processing Technology Office

23 July, 2007

Deep Green Supercharges the Commander



The Deep Green Vision

- Enhances commanders' option generation and decision processes
  - *Plan understanding* from multi-modal human sketching and speech
  - Intuitive presentation of decision points and second- and third-order effects of decisions
- Avoids "black swans" (unanticipated plan breakage) by predicting likely futures and building options before they are needed
- Generates a broad spectrum of possible futures with an instant-response, multi-resolution combat modeler
- Streamlined transition to the battle command community

"All current battle command efforts provide incremental improvements over existing capabilities. The only thing that provides something I don't have today is Deep Green."
BG Justice, Deputy Program Executive Officer for Command, Control and Communications Tactical (PEO C3T)

Gaps and Barriers

- Army written requirements for machine support for *commander-driven* planning and execution (italics are DARPA's):
  - Commanders require capabilities to enable more *rapid decision – action cycles, with much less effort* (TRADOC)
  - Capabilities to automatically present commanders with *relevant potential courses of action* (ARMY PEO C3T)
  - Enhanced cognitive *understanding of battlefield dynamics* (ARMY PEO C3T)
- Commanders find it difficult to:
  - Construct more than a small portfolio of hand-crafted plans
  - Foresee impending plan breakage
  - Generate branches and sequels during operations *before* plan breakage

- New technology is needed for machine induction of intuitively expressed plans
- Existing AI planning & monitoring systems focus on full automation and can be brittle
- Current generation of combat models are too slow and require significant manual intervention

Today, commanders generate three courses of action, and then choose *one* to execute. Limited contingency planning is conducted; when the *one* plan breaks, the commander becomes reactive.

Deep Green is Commander-Driven

Option Sketches, Voice, & Decisions

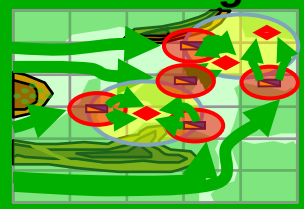
Animation, Cues, & Questions

Commander's Associate

Sketch to Plan

Sketch to Decide

Blitzkrieg

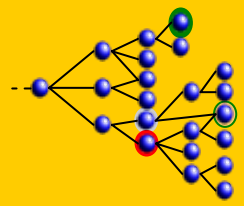


Candidate Plans

Possible Futures

Candidate Plans

Crystal Ball



Key Event & Decision Branch Points

Commander's Orders to Subordinates

Common Operating Picture

Environment
(Other Echelons, FBCB2, CPoF, PASS)

Deep Green is Commander-Driven

Option Sketches, Voice, & Decisions

Animation, Cues, & Questions

Commander's Associate

Sketch to Plan

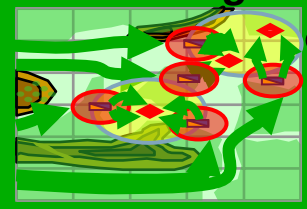
Sketch to Decide

Sketching and speech tools to help commanders generate options quickly

Intuitive, interactive graphical presentation to commanders of branching futures to help foresee downstream effects of decisions

Commander's Orders to Subordinates

Blitzkrieg



Fast, multi-resolution model to generate possible futures

- Near-OneSAF quality
- Highly accelerated playoffs of battlefield options and events

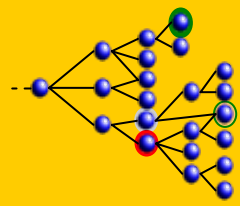
Candidate Plans

Possible Futures

Candidate Plans

Key Event & Decision Branch Points

Crystal Ball



- Maintain and evaluate diverse possible futures
- Proactively execute monitoring

Common Operating Picture

Environment
(Other Echelons, FBCB2, CPoF, PASS)

Rapid Transition to the Battle Command systems:

- Leverages CPoF software development kit and PEO C3T publish and subscribe system (PASS)

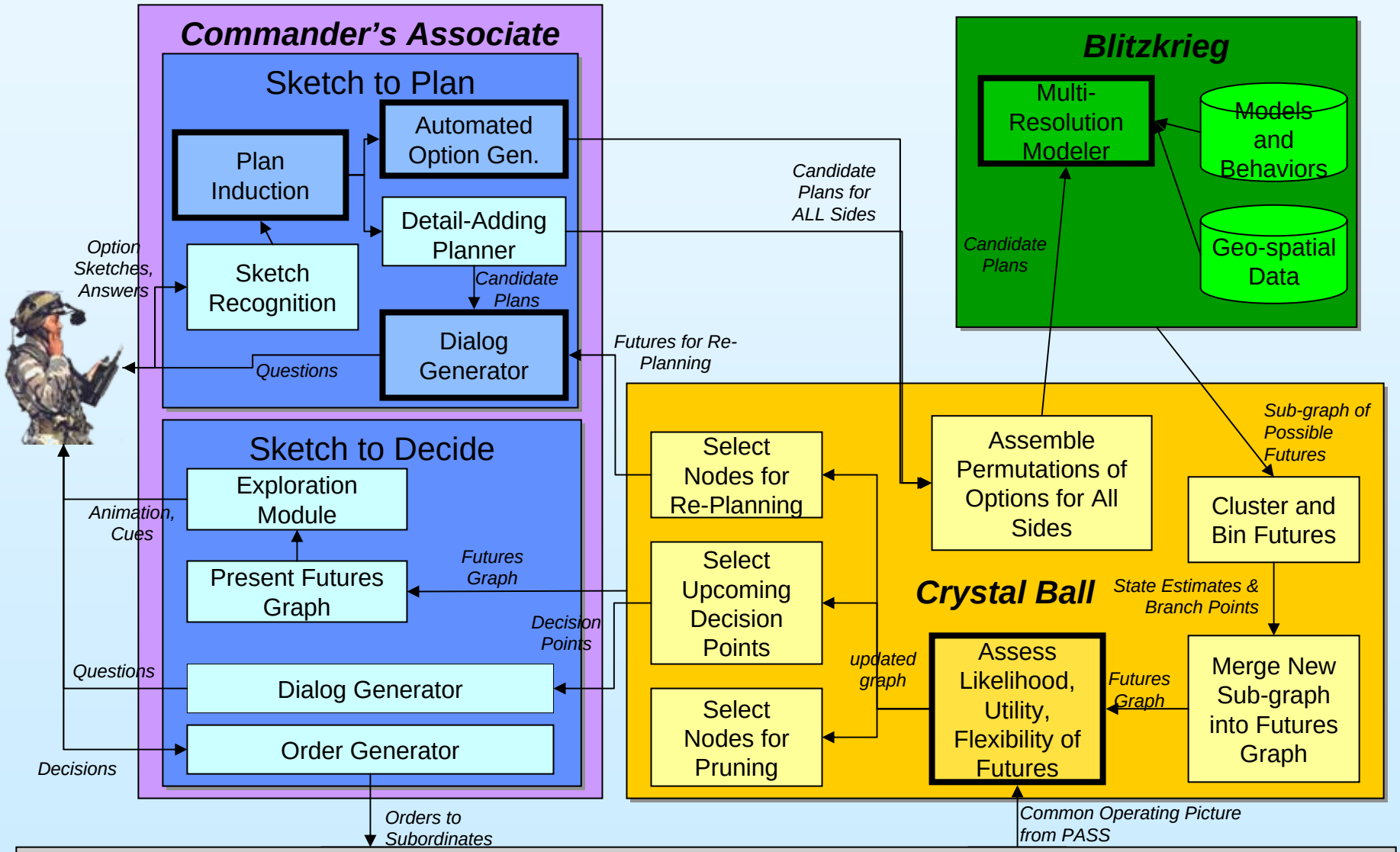
System Integrator sets and maintains data standards, performs component integration, transitions to battle command systems

Six Separate Tasks

- Task 1: Commander's Associate
  - 1a Sketch to Plan
  - 1b Sketch to Decide
- Task 2: Blitzkrieg
- Task 3: Crystal Ball
- Task 4: Automated Option Generation
- Task 5: Integration
- Task 6: Test and Evaluation

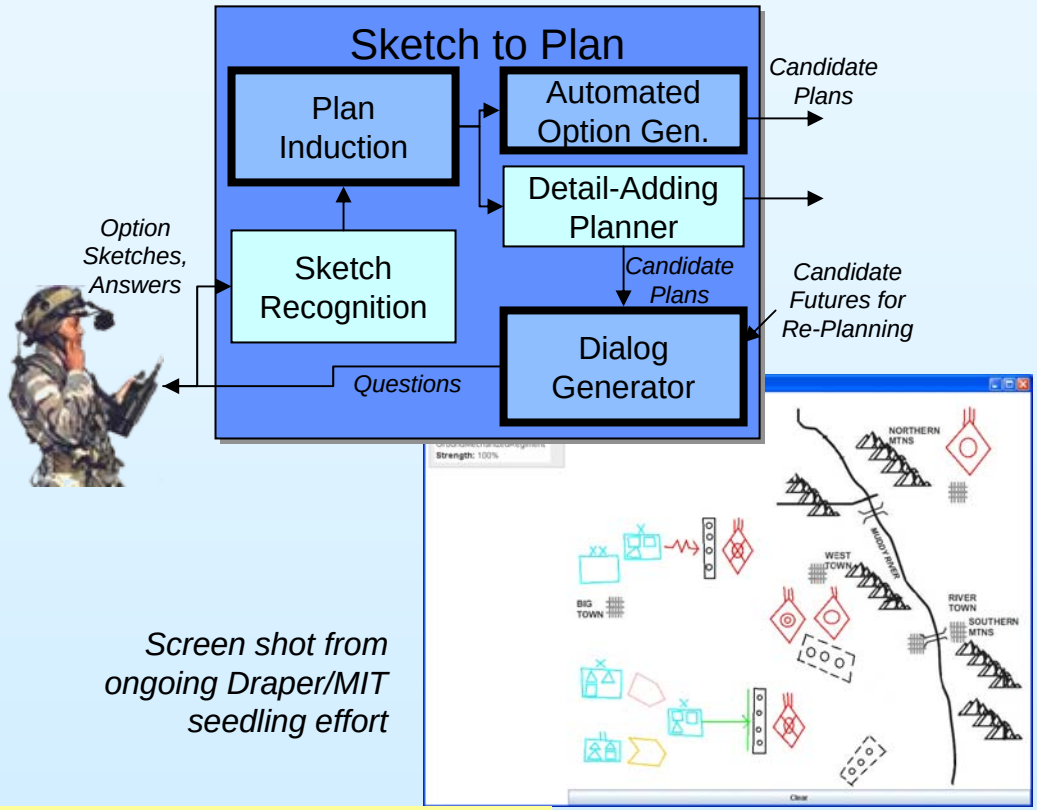
- Teams may bid multiple tasks, but if you are selected for test and evaluation, you cannot win any other contracts

More Detailed Architecture View



"Sketch to Plan" Concept

- Commander **draws** in "free hand" and **speaks**; STP interprets the symbols, replacing them with the correct standard military symbols
- STP *accurately* induces plans from sketch and speech, fills in missing details
- STP asks clarifying questions if it doesn't understand the sketch
- This allows commanders to specify an option at a coarse level, then move on to the next cognitive task



Screen shot from ongoing Draper/MIT seedling effort

DARPA Hard: Inferring plan from sketch understanding, including clarifying questions between the commander and **Sketch to Plan;** automated option generation

Why We Think We Can Get There:

- Current work in military sketching (Forbus, Cohen)
- Seedling effort with Draper Labs and MIT CSAIL (Davis)
- Current work on mixed-initiative planning (Allen, best AAAI paper 2007)
- Prior work on "natural" commander's intent language (Alphatech/BAE)
- Recent work in semi-automated plan generation (Myers, Wilkins), playbook approaches to semi-autonomous control (Goldman)

Task 1a: Sketch to Plan

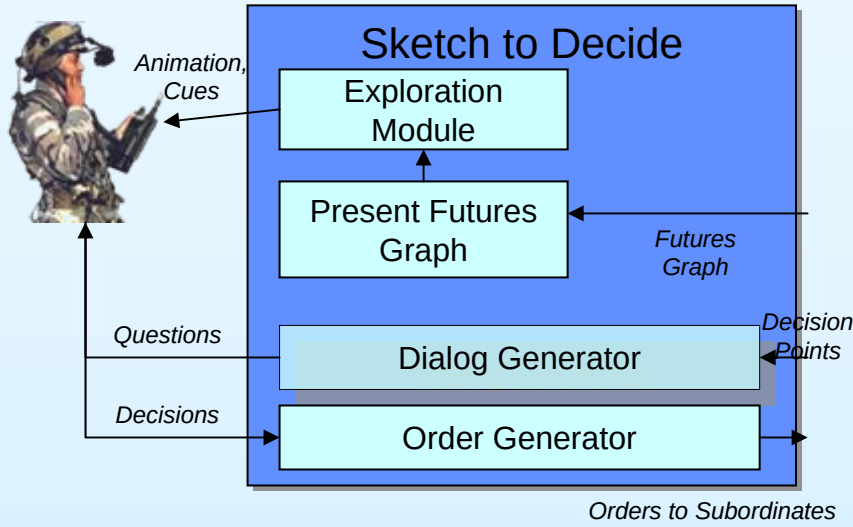
- Multi-modal (sketch and speech)
- Free-hand sketching, not pallet-driven
- Allow *COMMANDERS* and their staffs to generate options
- Detail adding planner assists, but does not generate options
- MIL STD 2525b recognition
- Provide user feedback
- Formulate clarifying questions
- Output: enhanced MSDL
- Expectation:
  - Technical approach to sketch understanding, plan induction, induction of commander's intent, and query generation

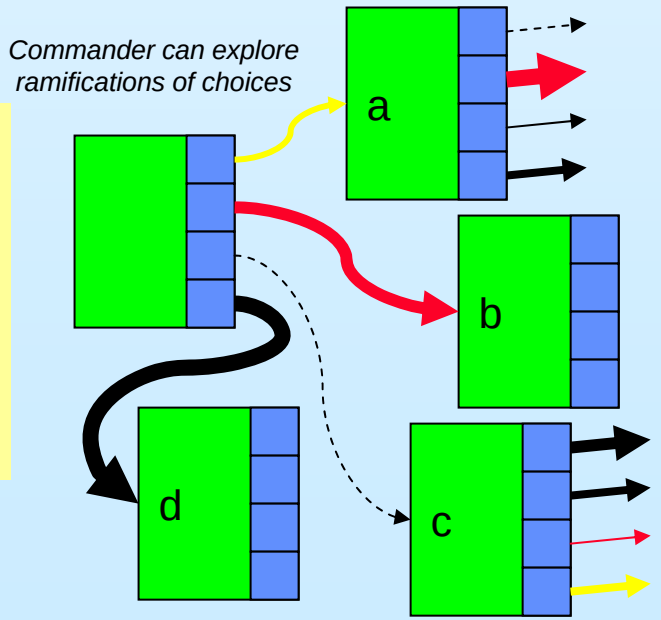
Task 1a: Sketch to Plan

- **Commander's Intent (Army):**
  - **A clear, concise statement of what the force must do and the conditions the force must meet to succeed with respect to the enemy, terrain, and desired end state. (FM 3-0)**
- **Commander's Intent (Marine Corps):**
  - **A commander's clear, concise articulation of the propose(s) behind one or more tasks assigned to a subordinate. It is one of the two parts of every mission statement which guides the exercise of initiative in the absence of instructions.**
- **The commander must not be constrained to specifying intent in artificial, stilted language; however, he should be constrained to use of doctrinally accurate terminology.**
  - **"The red zone" and "thunder run" are not donctrinal terms.**

"Sketch to Decide" Concept

- **Sketch to Decide** provides a window into the futures that have been generated by **Blitzkrieg**.
- Allows the commander to see how options may play out by following different "flows"
- Intuitive ways to depict likelihood, goodness, and flexibility as flows are explored





Hard: Presentation of branching possible futures to the commander in an intuitive way; Supporting user understanding of branch points across multiple decision dimensions and utilities

Why We Think We Can Get There:

•Recent efforts to generate "comic graphs" to show passage of time (Forbus)

•Recent work (Han) shows temporal relationships

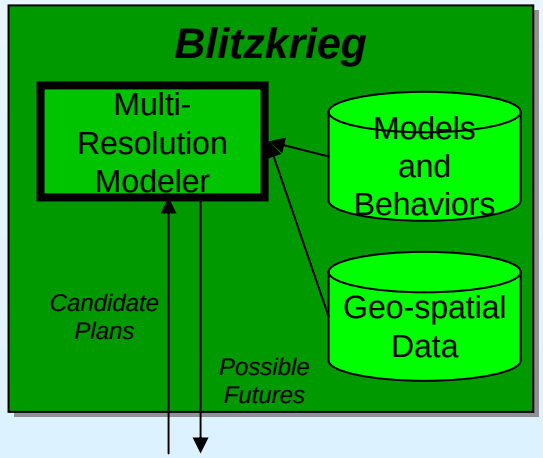
•Seedling effort with Draper Labs and MIT CSAIL (Davis)

Task 1b: Sketch to Decide

- Present multi-dimensional, decision-relevent information in a way that enhanced decision making
  - Enhance understanding of second- and third-order effects
- Exploration of the future graph
- From any “frame” of Sketch to Decide, the commander can perform Sketch to Plan
- Aids in commander’s visualization:
  - The mental process of achieving a clear understanding of the force’s current state with relation to the enemy and environment (situational understanding), and developing a desired end state which represents mission accomplishment and the key tasks that move the force from its current state to the end state (commander’s intent). (FM 6-0)
- Expectation:
  - Technical approach to presentation of multi-dimensional information that AIDS in understanding rather than increasing cognitive load

"Blitzkrieg" Concept

- High speed combat model links the portfolio of plans to potential battlefield outcomes
- Uses multi-resolution modeling (components of varying resolution) to achieve acceleration
- No human intervention during execution
- Uses hybrid of quantitative and qualitative methods to identify branch points
- Permits large time warping



DARPA Hard:

•Determining qualitative branch points and using quantitative techniques to assess likelihoods of each branch

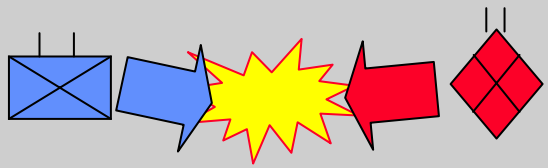
•Determining when and how to dynamically adjust model resolution to answer the appropriate question

•Massive time warping with little loss of fidelity

Why We Think We Can Get There:

•Multi-trajectory simulation research (Gilmer, Sullivan)

•Hybrid of quantitative modeling and qualitative physics (Forbus)



Example: two forces collide.
Qualitatively different outcomes include:

•Blue is destroyed
•Red is destroyed
•Blue begins to lose and withdraws
•Red begins to lose and withdraws
•The two forces choose not to engage
•And so on…

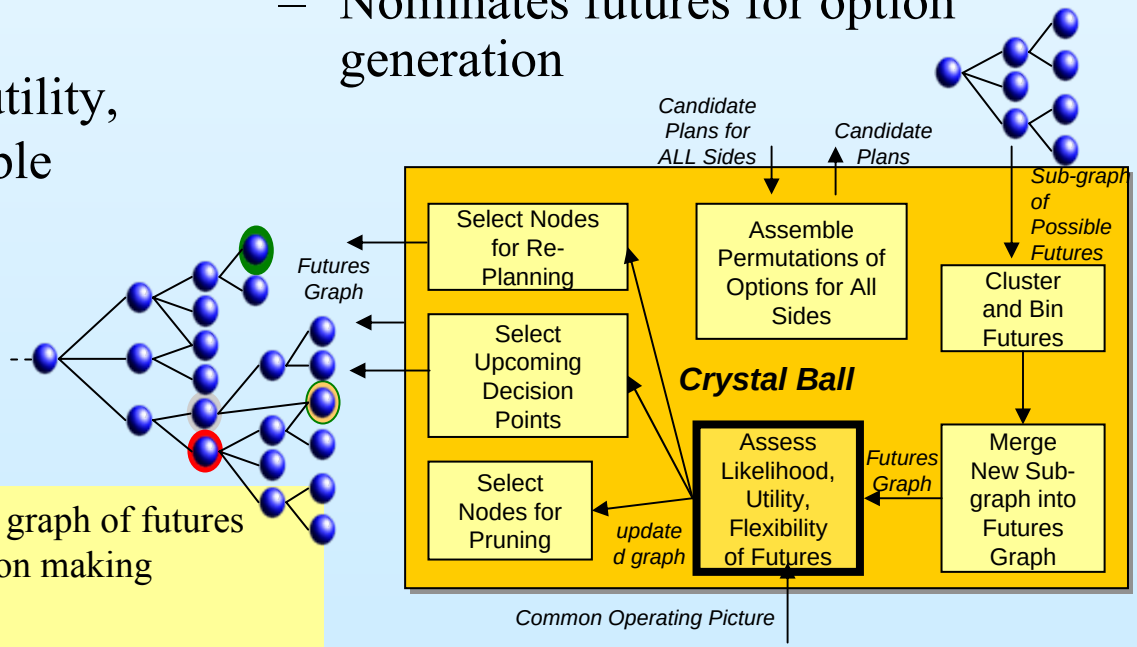
Task 2: Blitzkrieg

- Does not *do* planning; generates futures from options generated by humans
- Generates a broad set of qualitatively different futures, the “futures graph”
- May or may not have SAF-like intelligence of actors
- Probably aggregate level
- Integration of qualitative and quantitative methods
- Expectation: Offerors will describe the technical approach to building Blitzkrieg

"Crystal Ball" Concept

- During Option Generation
  - Clusters outputs of Blitzkrieg
  - Merges outputs from Blitzkrieg into graph of possible futures
  - Computes likelihood, utility, and flexibility of possible futures

- During Execution Monitoring
  - Monitors execution of operation
  - Updates estimates of likelihood, utility, and flexibility
  - Nominates futures for option generation



DARPA Hard: Maintaining an evolving graph of futures while identifying key branches for decision making

Why We Think We Can Get There:

•Seedling effort with ISI (Cohen)

•Recent developments in machine learning: Incremental case-based reasoning (Cox), Hybrid Bayesian networks (Pattipati), Structured Bayesian networks (Koller)

Task 3: Crystal Ball

- Does not *do* planning
- Collects information from battle command systems to assess the state of the current operation
- Maintains futures graph
- Predicts likelihood of possible futures
- Identifies futures for pruning, re-planning, and decision support
- Expectation:
  - Technical approach to maintaining graph
  - Details on how various metrics associated with futures will be computed
  - Transparency to facilitate explainability of results and recommendations

Task 4: Automated Options Gen.

- *Modest* effort
- Not the focus of Deep Green
- Given induction of plan and commander's intent, generate qualitatively different, relevant "mutations"
- Expectation:
  - Technical approach to creating qualitatively different mutations of the commander's plan

Task 5: Integration

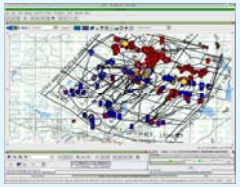
- Two main purposes
  - Interoperability among Deep Green components
  - Interoperability with real battle command systems
- Shouldn't write a lot of code
- *Independent*!!!
- Establishes data standards, interfaces, APIs, etc.
  - Expected to peer review recommendations
  - Enforce decisions
- Develop overall architecture in a way that supports the multi-echelon goal of Deep Green in subsequent phases

Task 6: Test and Evaluation

- BAA provides guidance on goals and metrics
- DARPA expects offerors to propose test plans and methodologies
- OneSAF Objective System as major part of test harness
  - OOS is ITAR protected!
- External support from ARL, HRED
- Within first 90 days
  - Provide test harness to component developers for at-home testing
  - Identify needed enhancements to OOS
  - Identify tentative test scenarios for Phases I and II
- Develop test scenarios
- Test readiness reviews
- Two tests per phase

Demonstration and Test Plan

Phase I
"Functional Components"
12 months



SAF Plan View Display

Deep Green

- Component tests
- Mid Intensity Conflict (e.g., 3rd ID during invasion of Iraq, Masr al Sharif)
- Test SMEs: Component testers
- SMEs interact directly with various components
- OneSAF used as exercise driver

Phase II
"Functional, Integrated System"
12 months

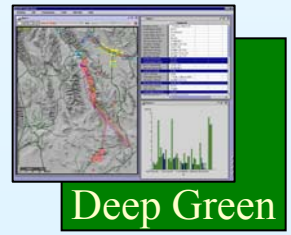
Battle Command

Deep Green

- A series of force-on-force, human-in-the-loop experiments in a simulation center
- Counter-insurgency operations
- Test SMEs: Commander, S-3 (operations officer), and S-2 (intelligence officer)
- Staff interacts directly with the Commander's Associate
- OneSAF used as exercise driver

Phase III
"Integrated with Battle Command"
12 months

Battle Command



CPoF with Deep Green "Under the Hood"

- A series of force-on-force, human-in-the-loop experiments in a simulation center and a tactical environment
- Large, Three-Block War Operation (e.g., Falluja or Basra)
- Test SMEs: Commander, S-3 (operations officer), fire support officer, engineer officer, and S-2 (intelligence officer)
- Staff interacts with Commander's Associate through CPoF
- OneSAF used as exercise driver

Go/No Go Criteria

Task	Phase I	Phase II	Phase III

Sketch to Plan: Accurate recognition of sketched MIL STD 2525b symbols.
F = (2 x H x P) / (H + P) x 100, where
#T = number of symbols that should be correctly interpreted
#I = number of symbols that are interpreted
H = Hit rate = number of correctly interpreted symbols / #T
P = Precision = 1 - False positive rate = number of correctly interpreted symbols / #I
	A subset of the MIL STD 2525b symbols (approx. 1/5th of the total) are commonly used when describing military operations. For this subset, F > 90%	For a subset consisting of 80% of all MIL STD 2525b symbols, F > 90%	For all MIL STD 2525b symbols, F > 90%

Sketch to Plan: Accurate machine induction of user's intended plan. Subject matter experts judge key aspects of machine-induced plans by playing them through OneSAF. S = (# aspects of plan judged to be correct / # aspects of plan entered by user) x 100	S > 70%	S > 80%	S > 90%
Crystal Ball: Reduce blind alleys during execution. A "blind alley" occurs during execution when Blue reaches an unpredicted or ill-prepared state (less than three good options available). A = (# blind alleys with Crystal Ball / # blind alleys without Crystal Ball) x 100	A < 80%	A < 50%	A < 20%
Blitzkrieg: Reduced time to evaluate combinations of representative Blue and Red courses of action. (Canonical test case will involve 3 Blue COAs vs. 3 Red COAs.)	Today: 120 min Deep Green < 30 min.	Today: 120 min. Deep Green < 10 min.	Today: 120 min. Deep Green < 3 min.
Overall System: Reduce staff requirements. Measured by reduction in staff usually needed to accomplish brigade planning/execution for the milestone scenario. T = # Staff With Deep Green / # Staff Without Deep Green	T<80%	T < 50%	T < 25%
Overall System: Commander's Performance. Measured by a single numerical score P, computed as a weighted function **f**(•) of mission accomplishment, friendly losses, enemy losses, neutral losses, time to accomplish tasks, and various human factors. P = **f**(with Deep Green) / **f**(without Deep Green)	P>= 1	P > 1	P > 2

General Notes (1 of 2)

- Tasks awarded separately, but this WILL be one team
  - Associate Contractor Agreements
- To be considered responsive, address all boxes from the architecture pictures for your components
- Propose Phase I with options for Phases II and III
- SMEs will be a program asset
- Planning for three 12-month phases
- Address all BFAs for each phase, as well as COE, etc. For instance, how can Crystal Ball facilitate more dynamic intelligence collection planning?

General Notes (2 of 2)

- Use services and functions allocated to other components of Deep Green rather than create redundant capabilities
- OOS is ITAR protected, so precautions have to be taken by ALL performers
- Reports from seedlings will be provided GFI to all performers
- All software, documentation, and data will be unlimited use
- IF there are more than one performer per task, down-selection will be made at the end of Phase I.
- All components as learning agents.